

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution printed.

CITIMENTS

Result No.	Query Score	Match length	DB ID	Description	
				DB	DB
c 1	24	120 0	102 17	A0010934	A021934 52367TF
c 2	24	100 0	106 9	A1003001	A1005001 an100509
c 3	24	130 0	109 17	A0825768	A0825768 2M01G
c 4	24	120 0	115 10	A0946335	AW46335 RC2-F10
c 5	24	120 0	127 17	A0761702	A0761702 1M056610
c 6	24	120 0	130 13	B0358857	BM42857 1454BC4
c 7	24	160 0	120 14	B0364004	B0364004 RC1-ST0
c 8	24	160 0	136 14	W86729	W86729 Eb63c11.r
c 9	24	120 0	137 12	HE774194	HE774194 MR1-OMO
c 10	24	100 0	138 17	B0111668	BH111668 SALK05
c 11	24	100 0	139 10	A0946443	AW946443 RC2-F10
c 12	24	100 0	143 17	BH101135	BH101135 RPC1-24
c 13	24	160 0	149 13	B0101135	BH101135 B1709837
c 14	24	160 0	151 10	B0063341	B0063341 BB06334
c 15	24	165 5	151 12	BG141368	BG141368 1a34905
c 16	24	100 0	151 14	B0113496	BH113496 ES16050
c 17	24	100 0	153 17	A0754818	A0754818 RNC1-23
c 18	24	100 0	155 17	AL764715	AL764715 Arabido
c 19	24	161 0	156 19	BB288429	BB288429 BB28842
c 20	24	100 0	161 9	AAB41154	AAB41154 MB31deAA64
c 21	24	100 0	162 10	BB288375	BB288375 BB28837
c 22	24	100 0	163 9	A0729893	A0729893 nx40c021
c 23	24	100 0	163 17	A0757569	A0757569 Arabido
c 24	24	100 0	166 12	BG141360	BG141366 1a33105
c 25	24	100 0	167 9	A0784858	A0784858 uc05105.
c 26	24	100 0	171 16	REF04-01*	REF04-01* ME1-EN0
c 27	24	100 0	173 10	B0288914	BH288914 BB28891
c 28	24	100 0	174 12	A0722321	A0722321 MR00356
c 29	24	100 0	177 9	AV164556	AV164556 RNC1-23
c 30	24	100 0	176 12	BG000266	RG000266 M2_GN0
c 31	24	100 0	176 12	BG000266	RG000266 M2_GN0
c 32	24	162 0	177 19	B0288394	B0288394 B328839
c 33	24	162 0	177 17	BH410918	BH410918 B028807
c 34	24	100 0	181 10	BH28845	BH28845 RC2848
c 35	24	100 0	181 12	BG000455	BG000455 RNC1-23
c 36	24	100 0	181 13	BG993173	BG993173 PM0-H10
c 37	24	100 0	181 14	BQ358496	BQ358496 CM1-HT0
c 38	24	100 0	184 17	A0258129	A0258129 GIT_HSF
c 39	24	100 0	185 9	A050393	A050393 5372105.
c 40	24	100 0	185 12	B059418	B059418 SV2_NNC
c 41	24	100 0	181 16	H0961211	H0961211 BBG621
c 42	24	100 0	182 9	AN626597	AN626597 ab54q07.
c 43	24	100 0	183 10	BB412668	BB412668 BB41266
c 44	24	100 0	193 10	B0503032	B0503032 bs78001
c 45	24	200 0	194 12	BH338602	BH338602 SM4_BH0

ATTACHMENT C

RESULT 1
AQ010934 /C
LOCUS AQ010934 102 bp **DN**A linear GSS 29-MAY-1998
DEFINITION F23C7RC :GF Arabidopsis thaliana genomic clone F23C7, DNA
SEQUENCE
ACCESSION AQ010934
VERSION AQ010934.1 GI:3165707
KEYWORDS CSS
SOURCE thale cress
ORGANISM Arabidopsis thaliana
EUKARYOTA: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
SPERMATOPHYTA: Magnoliophyta: Eudicots: Brassicales: Cole eudicots:
ROSIDAE: eurosids II: Brassicaceae: Arabidopsis
REFERENCE 1 (bases 1 to 102)
AUTHORS Boussov S.D., Sub F., Wibis S., Goldon K., Shatsman S., Chou D.

Yiu, K., Akininoto, H., Shen, K., Goelaskarath, S., Miltitsch, J., Adams, M. D. and Venter, J. C.	Seq primer: -40nt 3' fwd. RT from Amersham High quality sequence stop: 1.
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 4	FEATURES Source
JOURNAL Unpublished	Location/Qualifiers
COMMENT Author: OSSS: #2367FF Contact: Steve Kounsky Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: ronks@erigene.igb.ornl.gov Seq primer: M13 reverse Class: BAC ends High quality sequence stop: 102.	Location/Qualifiers
FEATURES Source	Site_1: SaliI; Site_2: EcoRI; Site_3: BpuRI; /site-Vector: Bcl1c111; Site_1: EcoRI; Site_2: BpuRI; /site-Vector: Bcl1c111; Site_1: EcoRI; Site_2: BpuRI; Produced by Thomas Attmann
BASE COUNT	24 a 16 q 36 t
ORIGIN	27 c
Alignment Scores:	Length: 106
pred. No.:	pred. No.: 1.27e+03
Score:	Length: 102
percent. Similarity:	Matches: 5
Best Local Similarity:	Conservative: 0
Query Match:	Mismatches: 0
DB:	Indels: 0
DB:	Gaps: 0
US-09-856-070-18 (1-5) x AG010934 (1-102)	US-09-856-070-18 (1-5) x A1003501 (1-106)
QY	QY 1 LysGluGluLeuMet 5
DB	DB 64 AAAAGAGAACCTGAG 50
RESULT 3	RESULT 3
LOCUS	AZ825768
ORGANISM	Mus musculus
DEFINITION	AZ010308F Mouse 13kb Plasmid JUGG1M Library Mus musculus genomic clone UGG2M0101G08 F, DNA sequence.
ACCESSION	AZ825768
VERSION	AZ825768.1 G1:29595676
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Fukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scutognathini; Muridae; Murinae; Mus. Islam, H., Aoyad, A., Barber, M., Beacroft, T., David, B., Ham, L. C., M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niedernhausen, A.
REFERENCE	1. (bases 1 to 109)
AUTHORS	Dunn, D., Aoyad, A., Barber, M., Beacroft, T., David, B., Ham, L. C., M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niedernhausen, A.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid insert.
JOURNAL	published (2000)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: diancgenetics.utah.edu
DEFINITION	Insert length: 10000 Std Error: 0.00 Plate: 0101 Row: G column: 08 Seq primer: RTGTTGAAATGAGCTGAT Class: plasmid ends High quality sequence stop: 109.
FEATURES	Location/Qualifiers
COMMENT	1. 109

Organism="Mus musculus"
 /strain="57R/6.1"
 /db_xref="taxon:100000"
 /clone-lib="Mouse 10kb plasmid library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tr1-resistant, F-"
 /note="vector: pMD20v, purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/phenotype/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA Polymerase and T4 polynucleotide kinase. Adapter oligonucleotides were ligated to the blunt ends in high molar excess. The adapter DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (9147321141qb) (AF139072.1), a copy-number inducible derivative of plasmid p1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptered mouse DNA was annealed to chemically competent E. coli XL1-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 29 a
 ORIGIN 34 9 31 t

Alignment Scores:
 Pred. No.: 1.4e+03
 Score: 24.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 17

US-09-856-070-18 (1-5) x AW946335 (1-109)

Qy 1 LysGluGluLeuMet 5
 Db 79 AAGGAAGGAACTAIG 93

RESULT 4
 AW946335 AW946335 115 bp mRNA linear EST 31 MAY 2000
 DEFINITION KC-2-ET0118-25640-012-e37 Erithrin Home sapiens cDNA, mRNA sequence.
 VERSION AW946335.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 115)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Patai,S., Costa,F.F., Brunstein,A., de Oliveira,P.S., Rucher,P., Tongneel,J.V., O'Hare,M.J., Soares,F., Krennani,R., Reis,L.F., de Souza,S.J., and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson,A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

BASE COUNT 38 a
 ORIGIN 20 c 24 9 33 t

Alignment Scores:
 Pred. No.: 1.51e-03
 Score: 24.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 16

US-09-856-070-18 (1-5) x AW946335 (1-115)

Qy 1 LysGluGluLeuMet 5
 Db 10 AAGGAAGGCTGATG 24

RESULT 5
 A2761701/c A2761701 127 bp RNA linear GSS 16-TEB-2001
 DEFINITION IM056707 Mouse 10kb plasmid genomic library Mus musculus genomic clone UGCC1M056J07 F, DNA sequence.
 ACCESSION A2761701
 VERSION A2761701.1
 GSS A2870903
 KEYWORDS SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 127)
 AUTHORS Iman H., Longacre,S., Mahoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Ticey,A., von Niederauern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 COMMENT Unpublished (2000)
 CONTACT Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddnagenetics.utah.edu
 Insert length: 10000 Std Error: 0.00
 Print: 0556 Row: J Column: 07
 Seq primer: CGTCGAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 127.
 FEATURES Location/Qualifiers 1..127
 Source /organism="Mus musculus"

This sequence was derived from the FAPESP/LIGR Human Cancer Genome project. This entry can be found in the following URL:
 (http://www.ludwig.org.br/series/series/0012_P112-312-002-ET0118_290490-012-0015-3209-04-29814-1)
 Seq primer: puc 18 forward
 High quality sequence start: 9
 High quality sequence stop: 115.
 Location/Qualifiers 1..115

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stae="Adult".

/note="Organ: uterus; Vector: puc18; Site: 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORFLESS PCR (U.S. Letters patent application No. 196
 ,716 - ludwig Institute for Cancer Research) Profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 26 a 32 c 29 g 50 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.94e-04 Length: 137
 Score: 24.00 Matches: 5
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0
 HS-09-856-070-18 (1-5) x BHBL1366 (1-138)

QY 1 LysGluGluLeuMet 5
 |||||||
 DB 67 AAGGAGGAGCTGAG 53

Alignment Scores:
 Pred. No.: 1.94e-04 Length: 137
 Score: 24.00 Matches: 5
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0
 HS-09-856-070 18 (1-5) x BE774194 (1-137)

QY 1 LysGluGluLeuMet 5
 |||||||
 DB 73 AAGGAGGAGCTGATG 59

RESULT 10
 BHBL1366 c BHBL1366 138 bp DNA linear CSS 02/MAY/2002
 /organism="Lycus" /definition="Arabidopsis thaliana genomic clone SALK_056499, DNA sequence."
 /source="thale cress, thale cress, Arabidopsis thaliana, Eukaryota, Viridiplantae, Streptophytida, Embryophytida, Tracheophytida, Spermatophytida, Magnoliophytida, eudicots, core eudicots, Rosidae, eudicots II, Brassicaceae; Arabidopsis".
 /reference="1 (bases 1 to 138)"
 /authors="Alonso, J. M., Leisse, T. J., Barajas, P., Chen, H., Cheuk, R., Gadrat, N., C., Jezek, A., Kernes, M., Kim, C. J., Parker, H., Prednis, L., Shirin, P., Zimmerman, J., and Ecker, J. R."
 /title="A Sequence-Indexed Library of Insertion Mutations in the Unpublished (2001)"
 /comment="Contact: Joseph R. Ecker
 Salk Institute for Biological Studies
 10010 N Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is similar to sequence recovered from the left border of TBLA. This sequence lies within an annotated exon of At3g17740 class: TBLA (aligned).
 FEATURES
 Source 1. 138 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="Colombia 0"
 /db_xref="taxon:3702"
 /clone="SALK_056499"
 /clone_id="Arabidopsis thaliana TBLA insertion lines"
 /note="This was performed on Arabidopsis thaliana lines each of which contains one or more TBLA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at <http://salk.salk.edu/~ecker/protocols.html>"
 BASE COUNT 40 a 47 c 25 g 31 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.94e-03 Length: 5
 Score: 24.00 Matches: 5
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0
 HS-09-856-070-18 (1-5) x BHBL1366 (1-138)

QY 1 LysGluGluLeuMet 5
 |||||||
 DB 67 AAGGAGGAGCTGAG 53

Alignment Scores:
 Pred. No.: 1.94e-03 Length: 5
 Score: 24.00 Matches: 5
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0
 HS-09-856-070 18 (1-5) x BE774194 (1-137)

QY 1 LysGluGluLeuMet 5
 |||||||
 DB 73 AAGGAGGAGCTGATG 59

RESULT 11
 AW946443 1 39 bp mRNA linear EST 31/MAY/2000
 /organism="Homo sapiens"
 /locus="S22-E10019-080509-012 f12 110319"
 /definiton="mRNA sequence."
 /accession="AW946443"
 /version="AW946443.1"
 /est="EST.61:8124212"
 /keywords="EST."
 /source="human."
 /organism="Homo sapiens"
 /note="Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo".
 /reference="1 (bases 1 to 139)"
 /authors="Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zaqi, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsuura, A., Baita, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Joncencel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J."
 /title="Shotgun sequencing of the human transcriptome with ORF expressed sequence tags"
 /journal="Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)"
 /comment="Contact: Simpson A. J. G.
 /laboratory="Laboratory of Cancer Genetics"
 /institution="Ludwig Institute for Cancer Research"
 /address="Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo, SP, Brazil"
 /fax="+55-11-27049222"
 /email="asimpson@ludwig.org.br"
 /tel="+55-11-27070001"
 /note="This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/~jz/est/est.html2/Est774194-08&t4-1>)
 /seq_primer="puc18 forward"
 /high_quality_sequence_start="13"
 /high_quality_sequence_stop="139".
 /location/qualities="1..139"
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="ET0019"
 /dev_stage="Adult"
 /note="Organism: Human. Vector: puc18. Site_1: small; Site_2: small. A mini-library was made by cloning products derived from ORFLESS PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 /base_count="48 a 24 c 31 g 36 t"
 /origin="ORIGIN"

Query Match:	100.0%	Indels:	0	RESULT 13
DB:	0	Gaps:	0	BJ008375
US-09-856-070-18 (1-5) x Aw946443 (1-139)				BJ008375
QY 1 LYSGIUGIuleuMet 5				BJ008375 MFOISSA cDNA oryzias latipes cDNA clone MFOISSA19c09 5', mRNA sequence.
DB 8 AAGGAACGCTGARG 22				BJ008375
RESULT 12				BJ008375 1 GL-17368489
BJ01135	143 bp	DNA	linear	149 bp mRNA linear EST 05-DEC-2001
LOCUS	RPCI-24-333H11	RPCI-24	Mus musculus	149 bp mRNA linear EST 05-DEC-2001
DEFINITION	DNA sequence.			
ACCESSION	BJ01135			
VERSION	GI:14925696			
KEYWORDS				
ORGANISM	Mus musculus			
SOURCE	house mouse.			
REVIEWER	1 (bases 1 to 143)			
AUTHORS	Zhao, S., Niemann, W., Malek, J., Shatsman, S., Aknnef, B., Levin, M., Isegawa, G., Geet, K., Krol, M., Shvartsbeyu, A., Gobidegeorgis, E., Russell, I.D., de Jong, P. and Fraser, C.M.			
TITLE	Mouse BAC End Sequences from Library RPCI-24			
JOURNAL	Unpublished (1999)			
COMMENT	Other GSS: RPCI-24-333H11.TJ			
CONTACT	Shayring Zhao			
DEPARTMENT	Department of Eukaryotic Genomics			
THE INSTITUTE FOR GENOMIC RESEARCH				
9712 Medical center dr , Rockville, MD 20850, USA				
TEL:	301 838 0200			
FAX:	301 838 0208			
EMAIL:	szhao@igc.org			
Clones are derived from the mouse BAC library RPCI-24. For: BAC				
library availability, please contact Pieter de Jong				
(pdj@jengmail.cho.org). Clones may be purchased from BAC/PAC				
Resources (http://www.chori.org/bacpac/orderinfoframe.htm) BAC end				
page: http://www.chori.org/bacpac/orderinfoframe.htm				
plate: 333 , row: H , column: 1				
Seq. primer: T7				
Class: BAC ends				
FEATURES	location/Qualifiers			
1. 143 bases				
Source	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="Taxon:10090"			
	/clone="RPCI-24-333H11"			
	/clone_id="RPCI-24"			
	/sex="Male"			
	/cell_type="Spleen/Brain"			
	/note="Vector: PTB2BAC1; Site_1: BamHI; Site_2: BamHI;			
	RPCI-24 Mouse BAC library produced by Pieter de Jong. The			
	library was cloned in the PTB2BAC cloning vector at the			
	BamHI sites using MboI. Partially digested male C57BL/6J			
BASE COUNT	34 a	33 c	47 q	29 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2.07e+03	Length:	143	
Score:	24.00	Matches:	5	
Percent Similarity:	100.0%	Conserv. lev.:	0	
Best Local Similarity:	100.0%	Mismatches:	0	
Query Match:	100.0%	Indels:	0	
DB:	17	Gaps:	0	
US-09-856-070-18 (1-5) x BH101135 (1-143)				
QY 1 LYSGIUGIuleuMet 5				
DB 60 AAGGAGGACTGATG 74				

Search completed: January 16, 2003, 21:37:05
Job time : 664.571 secs

